Bioestatística

Samuel de Oliveira Ribeiro

São Cristóvão/SE
2009

Bioestatística

Elaboração de Conteúdo
Samuel de Oliveira Ribeiro

Capa e Projeto Gráfico
Hermeson Alves de Menezes

Diagramação
Neverton Correia da Silva

FICHA CATALOGRÁFICA PRODUZIDA PELA BIBLIOTECA CENTRAL
UNIVERSIDADE FEDERAL DE SERGIPE

Ribeiro, Samuel de Oliveira.
R484 Métodos Quantitativos em Biologia I / Samuel de Oliveira Ribeiro -- São Cristóvão: Universidade Federal de Sergipe, CESAD, 2009.

1. Biometria. 2. Biologia 3. Título.

CDU 57.087.1

Sumário

Aula 1

CONCEITO E APLICAÇÃO DA ESTATÍSTICA, APRESENTAÇÃO TABULAR, DISTRIBUIÇÕES DE FREQUÊNCIAS E APRESENTAÇÃO GRÁFICA

META

Trabalhar a construção de Tabelas e Diagramas (gráficos) para todo tipo de Série Estatística.

OBJETIVOS

Ao final desta aula, o estudante deverá:
Elaborar tabelas sobre Distribuições de Freqüências por valores e intervalo de valores, bem como em relação a outras Séries Estatísticas.
Construir Gráficos para cada tipo de Distribuição de Freqüência ou Séries Estatísticas.

PRÉ-REQUISITO:

Ter em mãos: esquadros, compasso, calculadora, papel milimetrado e quadriculado.

INTRODUÇÃO

Olá! Tudo bem? Você está preparado para começar a trabalhar com números. Não é nada difícil muito pelo contrário é bastante fácil e levará você a manter contato com técnicas bastante interessantes. Com certeza vai ficar cativado pelas várias alternativas que este estudo pode oferecer.

A partir desta primeira lição vai entender com facilidade com gerar estatísticas de atividades corriqueiras de sua vida, bem como daquelas relacionadas com o bem estar econômico e social das pessoas que vivem em determinada comunidade, em um município, Estado ou País.

O importante é que você se interesse em trabalhar com números desde esta primeira aula. A aplicação da estatística no estudo de qualquer variável sempre envolve conhecimentos desta ciência que vão desde a coleta de informações até a análise e interpretação de resultados, portanto os temas tratados nesta primeira aula são fundamentais em qualquer pesquisa que o aluno venha a realizar.

Nesse primeiro momento você vai aprender como um Banco de Dados seja ele pequeno ou grande pode ser resumido em uma tabela de dados, sem nenhum prejuízo para o resultado da pesquisa.

Usando esta ferramenta de construção de tabelas, passamos a dispor de uma outra também muito importante que a elaboração de gráficos, facilitando deste modo à visualização e compreensão do comportamento das informações estatísticas mesmo por usuários não muito acostumados com pesquisas.

Tendo paciência e boa vontade em apreender e com certeza você vai ficar muito satisfeito com os resultados que vai conseguir.

ESTATÍSTICA BÁSICA

"É o estudo de fenômenos relativos a uma população, observando seu comportamento e identificando sua tendência para o futuro". É, sobretudo, um método de observação, seja através de pesquisa direta, seja através de comparação de fenômenos semelhantes, identificando suas características.

Portanto, podemos concluir que a Estatística não é simples coleta de dados, ela estuda a probabilidade de ocorrência do fenômeno ou estuda seu comportamento no tempo. A Estatística deixou de ser simples catalogação de dados numéricos sobre fenômenos coletivos, para ser o estudo de como chegar a conclusões sobre o universo a partir da observação de amostras.

De acordo com a forma de observações das populações a pesquisa estatística pode ser: **Descritiva** e **Indutiva**.

Descritiva - é aquela que parte de um conjunto de dados e obtém conclusões sobre os mesmos, não passando do conjunto de conhecimentos fornecidos por estes dados.

Indutiva - é aquela que ultrapassa os limites do conjunto de conhecimentos fornecidos pelos dados observados, isto é, a partir de uma amostra estuda as características da população amostrada.

FASES DO TRABALHO ESTATÍSTICO

- Planejamento, Coleta, Crítica, Apuração, Apresentação e Análise dos dados.
- Planejamento - é o detalhamento de todo o processo de pesquisa, de acordo com o objetivo do trabalho.
- Coleta - é o levantamento de dados sobre o fato a ser analisado. Pode ser Direta e Indireta.
- Direta **-** quando os dados são obtidos na fonte originária. (Questionários, cartas, entrevistas, observações, etc.).
- Indireta - quando os dados são copilados de outras fontes que não a originária. (Anuários, publicações, etc.)
- Críticas - é utilizada para avaliar a qualidade dos dados obtidos, rejeitando respostas falsas e. realizando possíveis correções para homogeneização dos dados.
- Apuração - Manual, Mecânica e Informatizada.
- Apresentação - Numérica e Gráfica.
- Análise - Estudo e conclusões sobre a população investigada.

APRESENTAÇÃO TABULAR

Uma tabela estatística compõe-se de elementos essenciais e complementares

ELEMENTOS ESSENCIAIS

Título - é a indicação que precede a tabela, contendo a designação do fato observado, local e época em que foi registrado.
Cabeçalho - é a parte superior da tabela que especifica o conteúdo das colunas.
Coluna Indicadora - é a parte da tabela que especifica o conteúdo das linhas.
Corpo - é o conjunto de colunas e linhas que contém, respectivamente, as informações sobre o fato observado.
Casa - é o cruzamento da coluna com a linha no corpo da tabela. Não existindo informações as casas devem ser preenchidas com algum sinal convencional, cujos prinscipais são os serguintes:

— Quando não existe a informação pesquisada.
... Quando não se dispõe dos dados. Ex: falta de informações sobre as precipitações pluviométricas em Aracaju, no mês de junho.
? Quando há dúvida sobre a exatidão do valor. É colocado a esquerda da informação.
& Quando o dado retifica informação anteriormente publicada. É colocada a esquerda da informação.
x Quando o dado for omitido a fim de evitar individualização da informação.

Elementos Complementares (situam-se no rodapé da tabela).

Fonte - é a indicação da entidade responsável pelo fornecimento dos dados ou por sua elaboração.
Nota - são informações de natureza geral, destinadas a conceituar ou esclarecer o conteúdo das tabelas ou indicar a metodologia adotada no levantamento ou na elaboração dos dados.
Chamadas - são informações de natureza específica sobre determinada parte da tabela. São identificadas por algarismos arábicos ou asteriscos, sempre a esquerda de cada informação e a direita da coluna indicadora, em ordem crescente de cima para baixo e da esquerda para a direita.

Unidades de Medidas - devem ser usadas sempre no singular e em letras minúsculas. Grama (g) Quilo (kg) Tonelada (t) Metro (m) Segundo (s) Minuto (min), etc.

Data de Referência dos Dados

- Anos civis consecutivos: 1960-70; 1890-910
- Anos civis não consecutivos: 1950-1965

- Período de doze meses diferente do ano civil: 1960/65

Principais normas para construção de tabelas

- As tabelas são delimitadas na parte superior e inferior, exclusive a esquerda e a direita.
- É Facultativo o emprego de traços para separar as colunas no corpo da tabela.
- Quando a tabela ocupar mais de uma página por excessiva altura, só será delimitada na última página. Por excessiva largura, divide-se a tabela em quantas partes forem necessárias. Em ambos os casos usam-se devidamente as expressões continua e conclusão.
- Quando uma tabela ocupar páginas confrontantes, todas as linhas devem ser numeradas na 1ª e última coluna. O disposto acima não sendo possível, deve-se desmembrar a tabela em seções uma embaixo da outra, separadas por um traço horizontal duplo.

SÉRIES ESTATÍSTICAS

É um conjunto de dados numéricos relativo a fatos observados, expressando as variações qualitativas ou quantitativas através de números por ordem de grandeza.

Elementos Fundamentais da Série Estatística

- Época (tempo) - período relativo a coleta de dados.
- Região (espaço) - local onde se passam os fatos observados
- Fenômeno (fato observado) - variável pesquisada

TIPOS DE SÉRIES ESTATÍSTICAS

Série Histórica, Evolutiva, Temporária ou Cronológica - quando a variável é o tempo, permanecendo fixos o local e o fato observado.
Série Geográfica, Territorial ou Espacial - quando a variável é o local, permanecendo fixos o tempo, e o fenômeno observado.
Série Específica ou Categórica -quando a variável é o fenômeno descrito, permanecendo fixo o tempo e o local.
Distribuição e Freqüência ou Série de Freqüência - são aquelas que permanecem fixos o tempo, o local e o fenômeno descrito, sendo que este é apresentado através de gradações que concentram os resultados das observações.

Estudo de Distribuições de Freqüências

- Variável Discreta - aquela que assume valores em pontos, isto é, está associada ao processo de contagem. Ex.: número de alunos, número de crianças, número de prédios, etc.
- Variável Contínua - é aquela que pode assumir qualquer valor entre dois pontos. Esta variável resulta normalmente em mensuração. Ex.: Estatura, Peso, Idade, etc.
- Dados Brutos - são aqueles que ainda não foram numericamente organizados. Ex: Índice Pluviométrico registrado num período de 42 dias, no município "X".

146	152	159	160	141	147	163	178	157	174	172
159	151	166	148	157	151	162	173	141	165	152
142	158	148	162	155	176	150	143	158	168	
154	179	153	145	170	174	168	172	150	164	

Rol - É a organização dos dados brutos, segundo sua grandeza. Índice Pluviométrico: (n=42)

141	145	148	151	154	158	160	164	168	173	178
141	146	150	152	155	158	162	165	170	174	179
142	147	150	152	157	159	162	166	172	174	
143	148	151	153	157	159	163	168	172	176	

Amplitude Total (A) - é a diferença entre o maior e o menor valor do Rol ou da Distribuição de Freqüência. Ex.: A=179 - 141 = 38
Intervalo de Classe (h) - é a diferença absoluta entre o limite inferior e superior de classes sucessivas. Corresponde ao intervalo em que o ROL foi subdividido.

Para populações ou Amostras com 200 ou mais unidades utiliza-se para cálculo do intervalo a fórmula empírica de Sturges:

$$h = \frac{A}{1 + 3{,}322 \log n} = \frac{38}{1 + 3{,}322 \log 42} \cong 5$$

Para pequenas amostras podemos utilizar a referida fórmula com prioridades, quando for possível, para intervalos múltiplos de 10, múltiplos de 5 e números pares.

Classes de Freqüências - são os grupamentos em que a série foi subdividida, de acordo com a dimensão do Intervalo de Classe. São constituídas pelos valores da variável x, que se enquadram entre seus extremos: Li — Ls; Li |— Ls; Li—| Ls; Li |—| Ls.

ÍNDICE PLUVIOMÉTRICO	FREQÜÊNCIA (f_i)
140 |—145	4
145 |—150	5
150 |—155	8
155 |—160	7
160 |—165	5
165 |—170	4
170 |—175	6
175 |—180	3
$\sum$	**42**

Ponto Médio (x_i) - são os valores eqüidistantes dos extremos de cada classe de freqüência. $x_i = \frac{L_i + h}{2}$

FREQÜÊNCIAS DE CLASSES

Freqüência Simples Absoluta (f_i) - é o número de vezes que se repete um determinado valor de uma variável no caso de distribuição de freqüências por valores, ou o número de valores de uma variável contida em cada classe de freqüência caso se trate de distribuição de freqüência por classes de valores.

Característica de f_i

a) $\sum f_i = f_1 + f_2 + ... + f_n = n$ (número de observações do rol).

b) $0 = f_i = n$

Freqüência Simples Relativa *(fr_i)* - é o número relativo de unidades em cada classe de freqüência. É obtida dividindo-se cada freqüência simples absoluta pela freqüência total.

Característica de fr_i

$$\sum fr_i = fr_1 + fr_2 + ... + fr_n \qquad \sum fr_i = \frac{f_1}{n} + \frac{f_2}{n} + ... + \frac{f_n}{n} \Rightarrow \sum fr_i = \frac{1}{n}.n = 1$$

Freqüência Acumulada de uma Classe - é a soma da freqüência simples (absoluta ou relativa) desta classe com as respectivas freqüências simples das classes anteriores.

Principais Características

$$F_1 = f_1 \quad e \quad F_n = \sum f_i = n \qquad Fr_1 = fr_1 \quad e \quad Fr_n = \sum f_i = 1$$

Freqüência Percentual - é o produto da freqüência relativa (simples ou acumulada) por 100

$$f_i\% = fr_i \times 100 \quad \Rightarrow \quad \sum f_i\% = 100\% \quad e \quad F_i\% = Fr_i \times 100$$

APRESENTAÇÃO GRÁFICA

A representação gráfica de uma série estatística tem por finalidade a ilustração e a visão mais rápida e geral do fato observado com suas características particulares.

Principais Diagramas

- Curvas
- Colunas ou Barras
- Setores
- Polar
- Histograma
- Polígono de Freqüência

Diagramas em Curvas: este tipo de gráfico presta-se para a representação de séries cujos dados se apresentam em função do tempo (séries cronológicas). O princípio geral do gráfico em curva é caracterizado pelo par (x, y) de coordenadas que podem ser representadas num sistema cartesiano. Determinados graficamente, todos os pontos da série, basta uni-los por segmentos de reta.

Consumo Residencial de Energia Elétrica no Município X/SE. 2000-07	
Ano	Consumo (1000 Mwh)
2000	125
2001	140
2002	190
2003	250
2004	235
2005	215
2006	250
2007	285

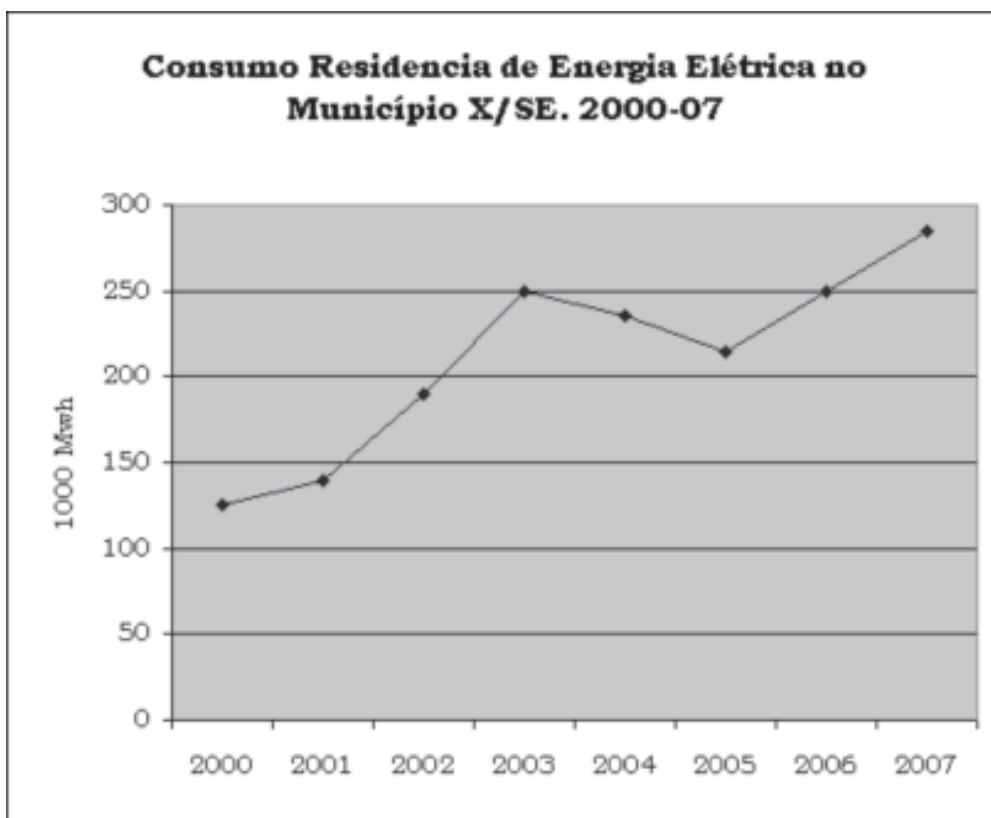



Diagrama Em Colunas - esta representação é feita por meio de retângulos com áreas proporcionais aos dados observados. A distância entre as colunas corresponde a 2/3 ou metade da base do retângulo. Sempre que

possível as colunas devem ser construídas em ordem crescente ou decrescente. O mesmo critério é utilizado para o diagrama em barras. Pode ser utilizados em séries históricas, geográficas e específicas.

Óbitos Menores de um ano da Região Nordeste, no ano X.	
Estado	Nº. de Óbitos
Piauí	5.800
Maranhão	5.350
Bahia	4.050
Pernambuco	3.715
Ceará	3.214
R/G do Norte	2.315
Alagoas	1.200
Paraíba	950
Sergipe	745
Total	27.339

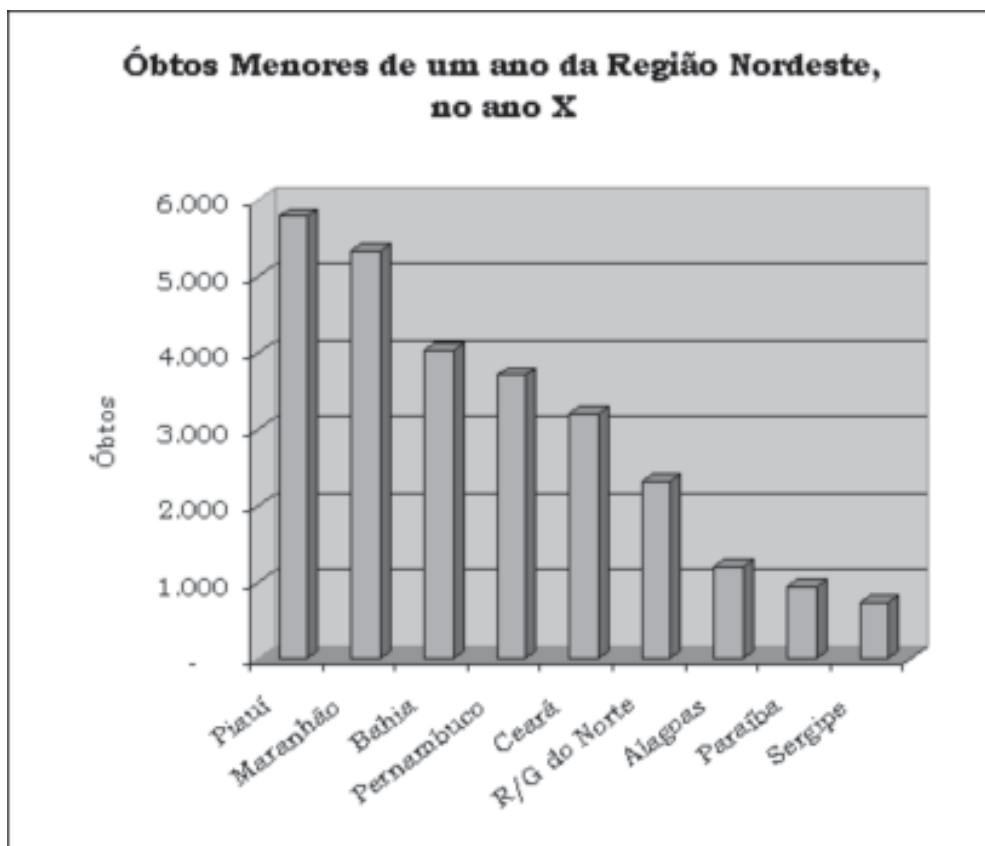



Diagrama Polar - é construído num círculo de raio arbitrário, dividido em partes iguais de acordo com o número de itens da série a ser representada. Marca-se, em cada raio vetor, o valor correspondente e, une-se os pontos por segmentos de reta. Tem grande aplicação na análise de séries históricas.

MOVIMENTO MENSAL DE COMPRAS DA EMPRESA DELTA 2005

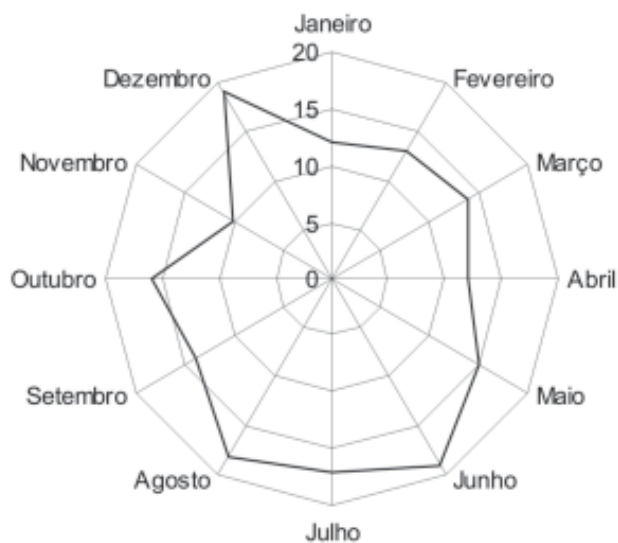
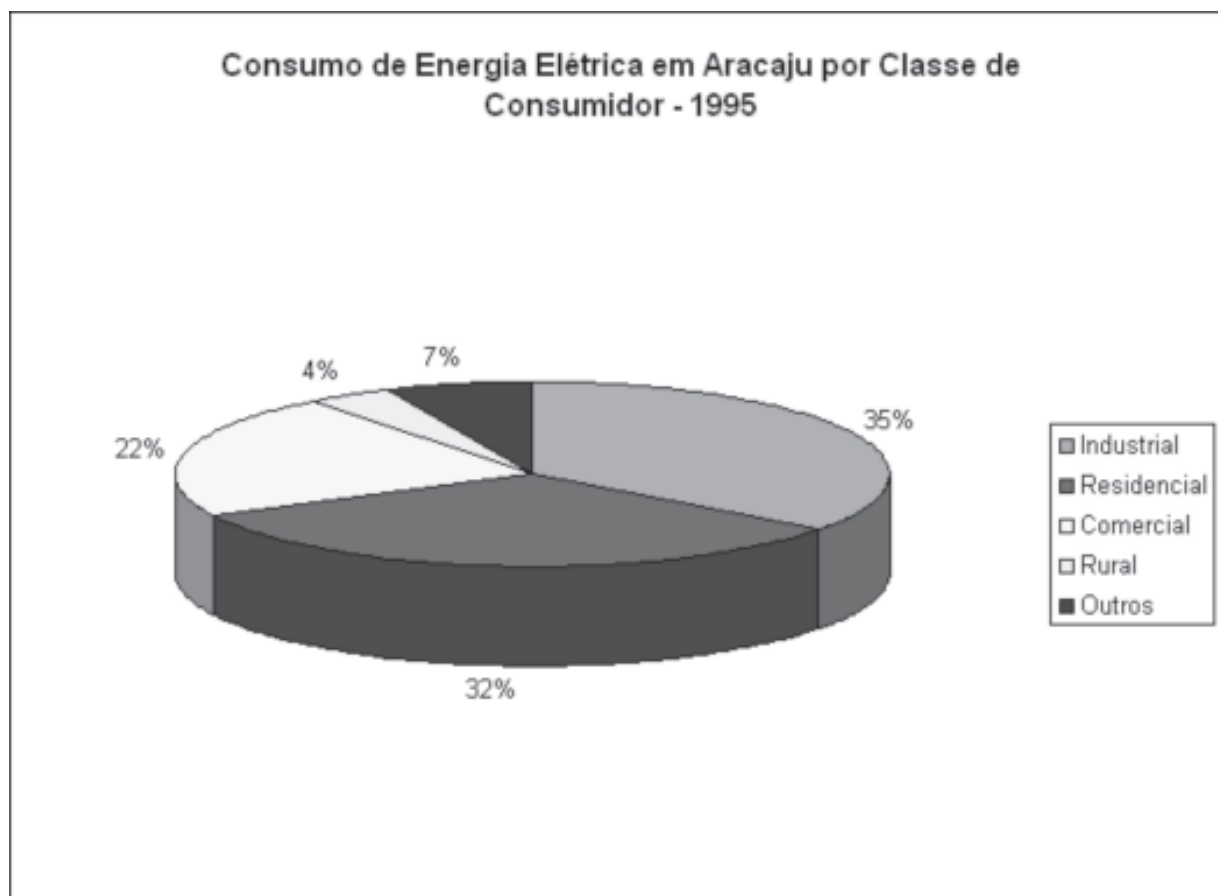
MESES	VALORES (R$1.000)
Janeiro	12
Fevereiro	13
Março	14
Abril	12
Maio	15
Junho	19
Julho	17
Agosto	18
Setembro	14
Outubro	16
Novembro	10
Dezembro	19

Cada ponto (P) do gráfico fica determinado pelo raio vetor e pelo ângulo polar (x, a). No gráfico Polar a coordenada angular (a) é constante, variando a coordenada linear x de acordo com o valor a ser representado.

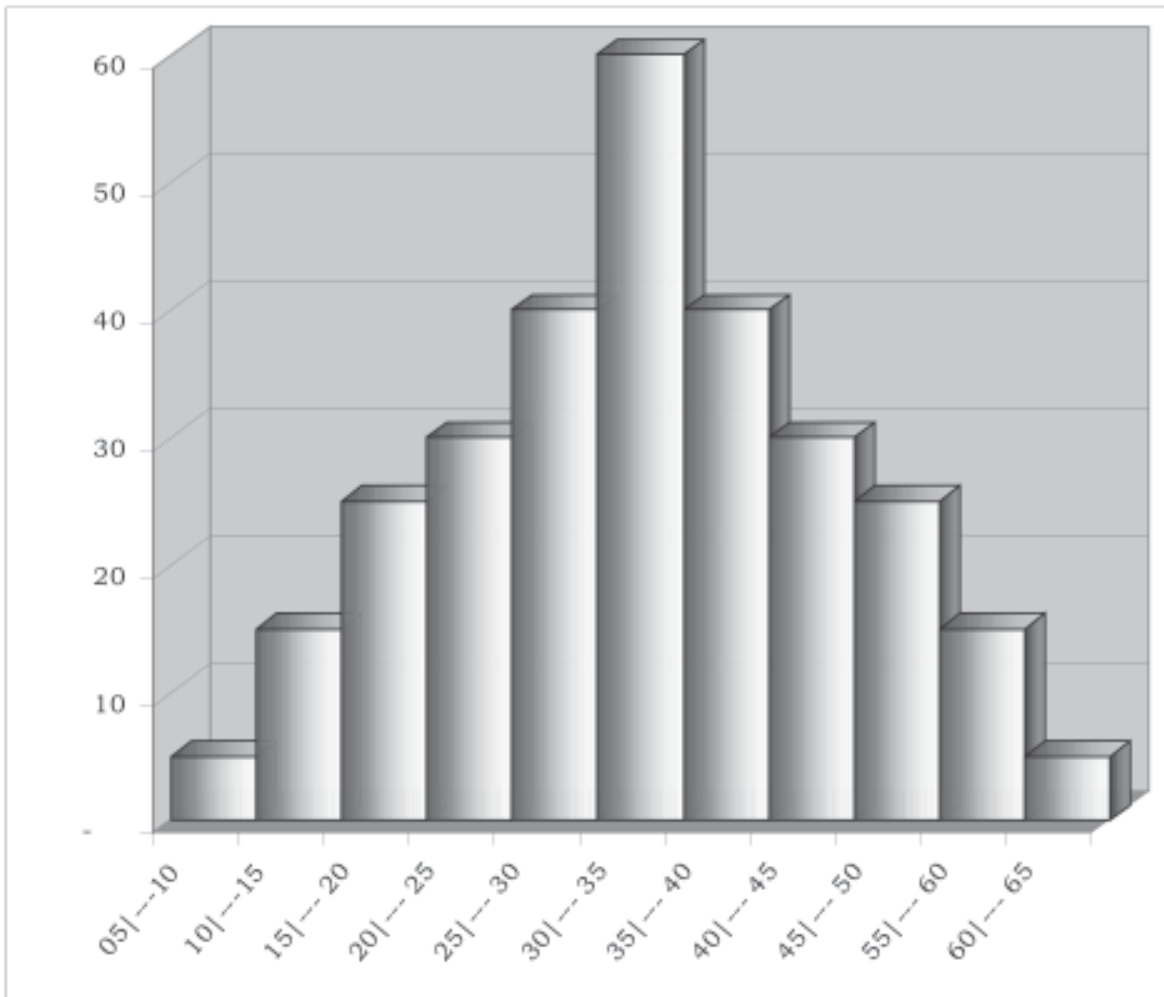
Diagrama Em Setores - é construído num círculo de raio qualquer, com ângulos centrais (setoriais) proporcionais as parcelas dos dados observados. Pode ser aplicado em séries geográficas, específicas e históricas.

Consumo de Energia Elétrica em Aracaju por Classe de Consumidor - 1995		
Classes	Consumo (1000 Mwh)	%
Industrial	225	35%
Residencial	205	32%
Comercial	140	22%
Rural	25	4%
Outros	42	7%
Total	637	100%

No diagrama devem ser registrado os percentuais dos setores, que são construídos em ordem decrescente no sentido horário, a partir do raio fixado acima do centro do círculo no sentido vertical.

Histograma - é um conjunto de retângulos adjacentes, com base no eixo das abcissas e áreas proporcionais às freqüências de classe.
Polígono de Freqüência - é constituído com base nos pontos médios das classes de freqüências da distribuição, ligados entre si por segmentos de retas.

Ex: Elaborar um Histograma para os dados abaixo:

Peso (kg)	Nº de Peças	Fi
05|—10	5	5
10|—15	15	20
15|— 20	25	45
20|— 25	30	75
25|— 30	40	115
30|— 35	60	175
35|— 40	40	215
40|— 45	30	245
45|— 50	25	270
55|— 60	15	285
60|— 65	5	290
Total	290	

60
50
40
30
20
10
05|---10 10|---15 15|--- 20 20|--- 25 25|--- 30 30|--- 35 35|--- 40 40|--- 45 45|--- 50 55|--- 60 60|--- 65

ATIVIDADES

1. Completar os dados que faltam para a seguinte distribuição:

X_i	f_i	F_i	fr_i
1	4	...	0,04
2	8	...	...
3	...	30	0,18
4	27	...	0,27
5	15	72	...
6	...	83	...
7	10	93	0,10
8	...	...	...
Σ			

2. Pontos obtidos por 100 alunos quando submetidos a um teste de conhecimentos. Elaborar a distribuição de freqüências e calcular: freqüência simples relativa, freqüências acumuladas e freqüência simples percentual. Elaborar o Histograma

15	30	10	40	43	40	28	32	14	18	40	41	47	10	35	10	40	19
27	45	23	50	35	45	47	45	42	29	33	23	30	45	20	40	20	50
34	23	40	47	28	10	15	24	50	47	24	28	35	10	34	40	47	40
38	40	29	50	14	23	38	40	19	47	16	40	30	40	28	26	28	45
30	25	33	40	28	28	15	19	20	17	35	14	10	19	24	15	40	34
35	28	12	40	25	22	35	25	40	29								

3. Elaborar um diagrama em colunas com os dados da tabela abaixo.

Renda Familiar das pessoas pesquisadas no Bairro Cidade Nova / Posto de Saúde José Quintiliano dez/07					
Meio de Transporte	Sexo				Total
	Masculino	%	Feminino	%	
Ônibus	5	42%	29	83%	34
Automóvel	3	25%	3	9%	6
Outros	4	33%	3	9%	7
Σ	12	100%	35	100%	47

4. Elaborar um diagrama em setores com os dados da tabela abaixo.

DISTRIBUIÇÃO DOS PACIENTES POR RAÇA - AGOSTO DE 2007		
RAÇA	Nº DE PACIENTES	%
Branca	146	41,71
Amarela	165	47,14
Negra	14	4,00
Não indicada	25	7,14
Total	350	100,00

5. Sejam os seguintes níveis de Glicose no sangue, em jejum, de 50 crianças: Obter uma distribuição de freqüência por classes e calcular as freqüências relativas (simples e acumulada).

51 61 57 63 60 57 82 68 59 72 68 73 75 66 61 76
72 75 64 69 68 69 55 65 67 59 66 76 65 64 71 72 73
54 81 60 80 66 52 71 75 74 73 65 74 84 56 62 83 62

CONCLUSÃO

Depois desta aula, você deve ser capaz de tabular qualquer conjunto de dados, tendo por base as Normas de Apresentação Tabular, bem como os principais elementos que devem acompanhar uma tabela. O mesmo se espera em relação à parte gráfica. Nesse caso o aluno deve saber como construir diagramas e tabelas feitos à mão, bem como utilizando o Excel na elaboração de Tabelas e Gráfico que com certeza irá facilitar bastante os estudos e análises estatísticas.

Alguns conceitos apresentados nesta aula podem parecer difíceis em um primeiro momento, podendo gerar em você dúvidas na elaboração de alguma tabela ou gráfico. Não deixe de recorrer a esta aula sempre que senti necessidade, vai encontrar nos temas aqui tratados uma base de informações bem detalhada e de fácil compreensão.

Com o passar do tempo você também vai perceber que o conhecimento aprendido nesta aula, vai facilitar o entendimento de outras aulas, bem como tornar as apresentações de resultados de pesquisas mais simples e com um maior poder de observação e análise.

Iniciando a segunda aula você vai perceber que tudo que apreendeu nesta primeira aula vai servir de base para aplicação do conteúdo a ser visto, assim sendo conjuntamente com esta aula você vai revisando e aplicando também os conhecimentos da aula anterior.

No final da aula temos uma lista de exercícios para serem resolvidos em grupos de no máximo cinco pessoas ou individual.

RESUMO

Nesta aula você aprendeu a trabalhar informações estatísticas oriundas de pesquisas ou de coletas indiretas.

No início deste aprendizado foram apresentados conceitos sobre Estatística, com detalhes iniciais para o estudo da Estatística Descritiva e

Indutiva, relativas à População e Amostra. Também foram apresentados os tipos de Séries Estatísticas e de Distribuições de Freqüências, com aplicação das principais Normas de apresentação tabular.

Depois desta fase introdutória passamos para a elaboração de tabelas, que a depender do tamanho da amostra ou população investigada pode ser apresentada em Distribuição de Freqüência por valores da variável pesquisada ou por intervalo de valores.

Na construção de tabelas foi indicado à técnica de cálculo de porcentagens para alguns tipos de freqüências, arredondamento de números e como utilizar os principais elementos de uma tabela.

A etapa final desta aula foi dedicada a Apresentação Gráfica, mostramos como construir os Diagramas mais utilizados na apresentação de dados estatísticos e como eles são importantes para melhor visualização e compreensão dos dados de uma tabela. Esta aula mostra passo a passo como estes diagramas devem ser construídos e para qual tipo de séries estatísticas eles são mais indicados. Estudando os diagramas você vai observar que existem algumas normas que não devem ser esquecidas como, por exemplo: Título, escala de valores, legenda etc. Além disto, os diagramas devem ser simples para facilitar o entendimento do usuário e ao mesmo tempo refletir com precisão as flutuações dos dados observados, visto que esta visualização no gráfico é muito mais fácil do que nas tabelas estatísticas, principalmente para o usuário não muito acostumado em observar ou avaliar flutuações numéricas.

AUTO-AVALIAÇÃO

Sou capaz de realizar cálculos como: porcentagens, arredondamento de números?

Sou capaz de entender a diferença entre Séries Estatísticas e Distribuições de Freqüências?

Sou capaz de elaborar Tabelas Estatísticas, bem como qualquer tipo de Diagrama?

PRÓXIMA AULA

Calculo de parâmetros médios e outras medidas de posição. Estes parâmetros podem ser calculados a partir de dados agrupados em rol ou em Distribuições de Freqüências.

REFERÊNCIAS

RODRIGUES, PEDRO CARVALHO. **Bioestatística**. Universidade Federal Fluminense.
FONSECA, JAIRO DA. **Curso de Estatística**. Editora Atlas.
OLIVEIRA, FRANCISCO ESTEVAM MARTINS DE. **Estatística e Probabilidade**. Editora Atlas.
TANAKA. **Elementos de Estatística**. Editora McGraw.Hill.
BARBETTA, PEDRO A. **Estatística Aplicada as Ciências Sociais**. Editora da UFSC.
GÓES, LUIZ A. C. **Estatística I e II**. Editora Saraiva.
DÍAZ, FRANCISCA; LOPES, FRANCISCO JAVIER. **Bioestatística**. Editora Thomson.